SEQ ID NO:1

```
   1  GGCAATGGGT CGAAATTCAT AGAATTTTGT GTGAGGTGCG TAGCGGCTCT
  51  GACAGGGGTG CTGCGCGGAG ATCTCTGGTC TCAGGTAGGG CGACAATGGA
 101  GAGGTGTTAG TTGCCCCCTG TATCGCTCTC TGCGTGGCGC ATTGGGTCAT
 151  CCTGCCCGGA CATATGATAT TCCGCTAGAG GATTACTGAT AGTTTCTGCC
 201  TGTCGGGCTT GTCGGGCTTG TCGGGCTTGT CGGGCTTGTC GGGCCTGTCC
 251  CTCTTGTCCC GCCTGTCCTC ACTTTTTCAC AATCAAAAAA TGGGCGAAGC
 301  CCTTCTTGTT CTATAGTTCT TATAGTTCAT ACGAAAATTA CACATAATTA
 351  TCAATAGCTT ATTCGCTTAA AAGGGAGTAA TTGGGCCGCA AAAGGGAGTA
 401  ATTGGGCCGC AAAAGGGAGT AATTGGGCCG CAAAAGGGAG TAATTGGGCC
 451  GATATCGGTT GTTTACATGG GGAGGAATCC CCTTAATCAT TTCTCCCCAT
 501  GGGAAAGACA ACACAAGTGG CCGCAGACCG GGCCTTCGAC CAGACAAAAA
 551  CTGTGCTCCC TGCCGAGGTG GCGAGAGGGG TCTATATGCG CAATCCGCCC
 601  CGCCTGCAGG CGCTCAAGCT CATGCATTTA ATGATAGCCA CTGCGGGCGG
 651  CCGCATGGCT GATGATGTGC GCCATGAAAT GCGGCTGGCC GACATTCGCG
 701  CAATCGACGG CATGAAAAAC CATGACCGTG AGAGCCTGAC CCCGCTGTTC
 751  GAGGAGCTAG CCGCTGCGGT GTTGACCCAT GATGACCCTG CAAAGATGAT
 801  CGTGACAGTC GGCGGCTTGG TCGATGAGGC GCGAATAGAC TACCGCCAGG
 851  AGGCAAGCGG CGAACTCCTA GTGACGTGGA CCTTCCGGAG TACATTCCGT
 901  CGTATGGCGG CGGAGTCGAA CCACTGGGCC ATTCTCGACC GTCAAACGGT
 951  ATTCCATCTC GGTAGTAAGT ATTCCGTGCT GCTGTTCCAG CACGTCTCTA
1001  GTCTCGCCAA TCTTGATCGG ATGAGCGCGA AAACCTTTAC GGTCCCCGAG
1051  TTGCGGGCGC TCCTTGGAGT GCCCGAGGGA AAGATGGTTC GTTGGAACGA
1101  CGTTAACAGA TTTGCTCTCA AACCTGCACT GGATGAGATC AACCATTTAT
1151  CGCGTCTGAC ATTGACGGCA AAGCCGACCA AGATTGGCCG TAGCGTGGCA
1201  AGTGTGACTA TAGGCTGGGA AGTGAAAGAC GACCCAACCG TCGCCAGGCG
1251  CGAGCTGGCG GGTTCCAAGG TCGGTCGAGA TGCTCGTCGC AGAGGGGCAG
1301  CGGAAACGAT AGCCCCCTCC TTCCCAGAAG CGGGCGGGAT CACCTACAGT
1351  CCACGTTGGC TGGAGCTGAA ACGCTCTGCT GGCAGCAACA AGGACAACGA
1401  TCTGATCGCC TCAGACTTCC GGCGTTTCTG TCGGGAGAGA GGCGTGCGTC
1451  TGGACGCTGC AAACATCGAA AAACTGTTTT TAGATTTCTG CGCAAAGGTA
1501  GGGAAGGTTT GAGTTTTGAG GTATTTCACC GCAATAGTGT TAAATGACTT
1551  TCGTGAAACG ATGTGCAATA TAGCGGTAAG ACTATGAAAT ACACGGCTGG
1601  ACAGGCTGCA AAAGCAACGG GTGTGGCGAC CGCAACCATC ACTCGGGCGC
1651  TAAAAAGCGG TAAAATTTCC GGTAAAAAAG ATGAATCTGG GGCATGGGTT
1701  ATAGATCCTG CAGAATTGCA CAGAGTGTTT CCTCCCATTT CAAAGAAATA
1751  CACCGAAACA CCTAACACGC AAGTATATGG TAAGCGTGAT GAAACACATG
1801  AAATGACCTC AGAAATCAGC GCATTAGAGC GTGAAGTTCG GACTTTACGC
1851  GATGCTTTAT CTGATGCCAG GGAGGATCGC GACAAATGGC GCGACATGGC
1901  CGAGCGTCTT TCAATTTCAT CACCGATGAG AGAGGAAGAC CGCCCCCCTC
1951  AAAAACAAAG ATGGTGGAAG ATATTCTGAT CCTGGGCTTC AGGAGCCTTG
```

FIGURE 1A

```
2001  CCTTTACTGG CGGAAAAACG CGATATTGAG GCACAGGCCC GCACTTTAGA
2051  GGCGGAAGCC TATAACGAGT ACCAAAACAC TAGAAGCCAG ATTGAGGAAA
2101  ATAGGGAACG TG
```

FIGURE 1B

.

```
   1  TGGTGAACGC ATTGGCTTGA TGTTTGAGAA AAGCGAAAAG ACCCGGCCAC
  51  AGTTGTGGGT AGAGCGTCGA TATGTGCAAG ACCTGATGCT TGCTGACATC
 101  GAACTCCGTG TCTACCTCGC ATCGTCGCTG TATCAGCCTG CTGCGGATGG
 151  CGGAAAGCCC GCCTATGGTC GTCACGCAGC CCTTAAGGCG ATGCGCGACT
 201  TGGCCCATGC CGATCTGGTG CGTTTCACCA TCGGCCGGAT TACGCAACTG
 251  GAGATGATCC TAGAGCGGTT AACCGAGACA TCTGGTTAAC GCCATAAAGG
 301  CTGCGGCATG AAAATAGGCG GACAATCTGC GCTTGGCCGC CCCCGTTCTC
 351  AGCCGTGCTT GCTCTCTGCC TGCATGGCAC GACGCAGGAT CGCGTTCATA
 401  CGGGTCTGAT ATCCAGACCC GCCCGCCTTG AGCCATGCCA GCACATCGGC
 451  ATCAAGCCGC GCGGTGATCT GCTGCTTGAT CGGGCGATAG AAGCGCCCAC
 501  GCTCGGCGTC TGCCCATTGG GCTTCGGTCA GCTCGGGAAC ATCGTTGGTG
 551  TCGATCTGCT CGGGCGGCAG AGCGTCCAGC CGCGCCAATT TCTTGCGGCG
 601  CTCCTCGGTA AGAGCGGGCA GCGTATCGAA GGTGTATTCA ACCATTGGCA
 651  TATCTCTTCC TTTCCTGCGG TGTAGCGCGG CGAGCCGAAA TGATGCGGAT
 701  CGTCTCGACC GGATCGGGGC CAGCCTCGAT GATCAGGTGG GCAACCAGAA
 751  GGACGGCAGC GCCATAGATC TGCCCAACGG TTTGCCAGCG GTATTCCCCG
 801  CCCTCGATCC TATCCTGAAC CGTCAGGTGC AACGGATCGG CGAACACATG
 851  CACAGCATCC TCGAACCGGA TGCCATGCTT CTTTTCGTTC GTTTCCGCCT
 901  TGGCGGGATC CCAGATAAAC CGCATCTTCA TGGCAGAATT ATAACTACAC
 951  ATTTGTAGTT ATTCAATGGC AAGTCGCAGG TTCAAATCAC GCCCCCAAAC
1001  CGCAACTGTA TTCGTTCTAC TCACGCGCGC TTTTGAATAG AAGCTTGCAT
1051  GATAACACCC GCCGCGTCCT CAACAAAATA AGGCAAATCC GCCGCGCTGG
1101  CGCAATCTGC GCTTTGTCGA TGCAAGGTCT TGTGGTTTCA TACTGCAAGA
1151  GCATGCAAGG AATTGCCCCG GATGAGCACC ACGACGACAC CCACCAAGCC
1201  GGCCTGGAAC AAGGGCCGCG TTGTCGGGAA AAAGCCGCCG CTGACACCTG
1251  ACCAGATTGC CCTGATCCGT CTCATCCTGC GCCAGGAACG GGCGTGGCGG
1301  GATCTGGCTC TGTTCAACGT GGCGATCGAC ACCAGTTTGC GCGGCTCGGA
1351  CCTCGTGCGC CTGCGCGTCT CGGATGTGGC GACCCCAGCT GGTCTGCGTG
1401  AGATCGTCGA GATCCGCCAG AAGAAGACCG AGGCCCGCAA TGTCCGCCCC
1451  GTACAGGCCC GCCTGTCGGA GGGGACACGC GAGAGCCTGC GGGTCTATCT
1501  CGCGGCCTCT GACAAGCCGC TGCACAGCTG GCTGTTCACC GGACAGGGCA
1551  TCCGCTGGTC CCACACCCAC CTTAGCGAGA GCCAGCTGTG GCGCCTGTTC
1601  AAGTCCTGGC TCGAGAAGGC GCGGCTCGAT CCCAGCCTCT ACGGGCTGCA
1651  CTCGCTGCGC CGAACCTTCC CCAGCCACAT CTACCGCGAG ACCGGCAATC
1701  TGCGCGCCGC ACAGCTGCTG CTGGGCCATG CCAGCATCGA GAGCACCAAG
1751  GAGTACATCG GCACCGAGCA AGCCGAGGCC CTCGATATCG CACGGAGGTA
1801  TCACCTCTAA CCCATGGAGA CCTATCTCGA GAAGCGCATC CCCGCCAAGA
1851  ACACAGCACG GTTCTACCGC ATGGCGGTCC TGCCGAACCT GTTCGGGGAA
1901  TGGACGACGT ATCGAGAATG GGGCCGCATT GGCATCAGCG GCCGCATCCG
1951  GCTCGATTGG TTTGAGAGTG AACAAGATGC CATCGCTGCG ATGCTCGCCA
```

FIGURE 2A

```
2001  TCGAGACCGC CAAGCGTCAG CGCGGGTATT GGCTCGAGCC CATCCAGATT
2051  GACATGTTCC CAGGGGCATA ACAGGCCATC AATGTAAGAG TGCAAGCGGA
2101  GCAAGCAAAA GCCATTTCAC AGTGAGGTGG CAGATGTTCC TGTTTCACAG
2151  TGAAAGCGCT GATGCTGTTT CCACGCCACA GACTGATACG ACCAAAGCAA
2201  CGGGGTCTGC CGCCACAGAC CGGTTCGCCG GCCACCCGCA GAAACGCAGG
2251  TAAAATGGCG ATTTCCGCAA AAAAACCGTG CAAATGATGG CAAATCACCA
2301  TCCAGTTTCA TCCTGAAACC CGTCGCTCAA CATGAACGAG CAGGCCATCA
2351  TCCAAGCCCC AGAAACGCGG TGCGGCGACT ACAGATGAGC GATGTTCTGG
2401  CTCATAGGCT GCAAGGCCCT GCAACAGTGA TTTCACCGTG AGATTGCAGG
2451  GTCTTTTGGC TCTCCCGCAA GAGCCACCTC AGGGTGAGCG AGCTAGCCGT
2501  CTAGGTTCAC AGTGAAATCG CTGAGGAGCG TTGCGGGGCT TATGGTTTGG
2551  CTGGTCACGT TGGCCATCGG AATGGAGCAT ACGATGGCTT CTACGCAGTC
2601  GAATCCTGAG GCTTCACGTG GGAAAAATAC GCTCCAAAAA AGCCCTGACC
2651  AAATCTTGGA AAAATTGCTT GAAAAGTTTG CTTCTAAAAA ACTGGGAACG
2701  AGATATGCAC GAGATCCCTT ACGAGTGCTG TAGGAGTAAT GCAGTGGACA
2751  AAAACGCCAT TTTTTGCCCC AGTAGGAGTA ATGGAGTGGT TATTTTTTGG
2801  GAGATTTTGC TTCAGTAGGA GTAACGCGTT GGTTAAATTT GCTTGATTGG
2851  CGGTTCAAAT CGACCACCGA GCTGCCGTTG GTCGTATTCG ATCTGCCCCG
2901  CAATTGGGCA CTTGCAGGCC ATCCCCCTGA ACTTCTGGCG ATGACCATTT
2951  CGAAGGCAAT GGGTCGAAAT TCATAGAATT TTGTGTGAGG TGCGTAGCGG
3001  CTCTGACAGG GGTGCTGCGC GGAGATCTCT GGTCTCAGGT AGGGCGACAA
3051  TGGAGAGGTG TTAGTTGCCC CCTGTATCGC TCTCTGCGTG GCGCATTGGG
3101  TCATCCTGCC CGGACATATG ATATTCCGCT AGAGGATTAC TGATAGTTTC
3151  TGCCTGTCGG GCTTGTCGGG CTTGTCGGGC TTGTCGGGCT TGTCGGGCCT
3201  GTCCCTCTTG TCCCGCCTGT CCTCACTTTT TCACAATCAA AAAATGGGCG
3251  AAGCCCTTCT TGTTCTATAG TTCTTATAGT TCATACGAAA ATTACACATA
3301  ATTATCAATA GCTTATTCGC TTAAAAGGGA GTAATTGGGC CGCAAAAGGG
3351  AGTAATTGGG CCGCAAAAGG GAGTAATTGG GCCGCAAAAG GGAGTAATTG
3401  GGCCGATATC GGTTGTTTAC ATGGGGAGGA ATCCCCTTAA TCATTTCTCC
3451  CCATGGGAAA GACAACACAA GTGGCCGCAG ACCGGGCCTT CGACCAGACA
3501  AAAACTGTGC TCCCTGCCGA GGTGGCGAGA GGGGTCTATA TGCGCAATCC
3551  GCCCCGCCTG CAGGCGCTCA AGCTCATGCA TTTAATGATA GCCACTGCGG
3601  GCGGCCGCAT GGCTGATGAT GTGCGCCATG AAATGCGGCT GGCCGACATT
3651  CGCGCAATCG ACGGCATGAA AAACCATGAC CGTGAGAGCC TGACCCCGCT
3701  GTTCGAGGAG CTAGCCGCTG CGGTGTTGAC CCATGATGAC CCTGCAAAGA
3751  TGATCGTGAC AGTCGGCGGC TTGGTCGATG AGGCGCGAAT AGACTACCGC
3801  CAGGAGGCAA GCGGCGAACT CCTAGTGACG TGGACCTTCC GGAGTACATT
3851  CCGTCGTATG GCGGCGGAGT CGAACCACTG GGCCATTCTC GACCGTCAAA
3901  CGGTATTCCA TCTCGGTAGT AAGTATTCCG TGCTGCTGTT CCAGCACGTC
3951  TCTAGTCTCG CCAATCTTGA TCGGATGAGC GCGAAAACCT TTACGGTCCC
4001  CGAGTTGCGG GCGCTCCTTG GAGTGCCCGA GGGAAAGATG GTTCGTTGGA
4051  ACGACGTTAA CAGATTTGCT CTCAAACCTG CACTGGATGA GATCAACCAT
```

FIGURE 2B

```
4101  TTATCGCGTC  TGACATTGAC  GGCAAAGCCG  ACCAAGATTG  GCCGTAGCGT
4151  GGCAAGTGTG  ACTATAGGCT  GGGAAGTGAA  AGACGACCCA  ACCGTCGCCA
4201  GGCGCGAGCT  GGCGGGTTCC  AAGGTCGGTC  GAGATGCTCG  TCGCAGAGGG
4251  GCAGCGGAAA  CGATAGCCCC  CTCCTTCCCA  GAAGCGGGCG  GGATCACCTA
4301  CAGTCCACGT  TGGCTGGAGC  TGAAACGCTC  TGCTGGCAGC  AACAAGGACA
4351  ACGATCTGAT  CGCCTCAGAC  TTCCGGCGTT  TCTGTCGGGA  GAGAGGCGTG
4401  CGTCTGGACG  CTGCAAACAT  CGAAAAACTG  TTTTTAGATT  TCTGCGCAAA
4451  GGTAGGGAAG  GTTTGAGTTT  TGAGGTATTT  CACCGCAATA  GTGTTAAATG
4501  ACTTTCGTGA  AACGATGTGC  AATATAGCGG  TAAGACTATG  AAATACACGG
4551  CTGGACAGGC  TGCAAAAGCA  ACGGGTGTGG  CGACCGCAAC  CATCACTCGG
4601  GCGCTAAAAA  GCGGTAAAAT  TTCCGGTAAA  AAAGATGAAT  CTGGGGCATG
4651  GGTTATAGAT  CCTGCAGAAT  TGCACAGAGT  GTTTCCTCCC  ATTTCAAAGA
4701  AATACACCGA  AACACCTAAC  ACGCAAGTAT  ATGGTAAGCG  TGATGAAACA
4751  CATGAAATGA  CCTCAGAAAT  CAGCGCATTA  GAGCGTGAAG  TTCGGACTTT
4801  ACGCGATGCT  TTATCTGATG  CCAGGGAGGA  TCGCGACAAA  TGGCGCGACA
4851  TGGCCGACG  TCTTTCAATT  TCATCACCGA  TGAGAGAGGA  AGACCGCCCC
4901  CCTCAAAAAC  AAAGATGGTG  GAAGATATTC  TGATCCTGGG  CTTCAGGAGC
4951  CTTGCCTTTA  AAACCTGAAT  CAGCATTCTA  GCGATGCTGA  TAAGAAGTAA
5001  ATATAGCCAC  AATAGAGCGG  CCATTTTCCA  TTCACATACA  GCTCATCATG
5051  TGATCAATAT  CAAGTATTGA  TATTCATCAA  TGGAGAAGAA  TTTACATGTA
5101  TCACAGGATC  ATCACAGCAT  TTGTTTTTGT  ATTTCTAAGT  GCTAACATAA
5151  CTATCGCTGG  CCCTAAAGAA  GATTGTACTA  TTGCAGTATC  TCACCTTGGG
5201  TTTCAGACCG  ATAATTACAG  CTTTGTCGAA  GCCGGTTTTT  TTGCCAGAGA
5251  GAGACACGTT  TTTGATGGTG  TAATAAACTG  CTACGTATCT  CATGATGGTA
5301  ACATCACAG  CATCATCCGG  GGCAACACAC  CTCTTATGGA  AGATGGATAT
5351  TATGGCCCAG  AAGTACTGGC  GGAAAAACGC  GATATTGAGG  CACAGGCCCG
5401  CACTTTAGAG  GCGGAAGCCT  ATAACGAGTA  CCAAAACACT  AGAAGCCAGA
5451  TTGAGGAAAA  TAGGGAACGT  GCCCTCGAGG  CGCTGCGGCT  AGCTAGCAGT
5501  CCTTTTATTA  ATAATGGTAG  TACAGAAGAA  CAGACAATTA  TACAGGCCGC
5551  AACTCCGACG  GCAGATCCTG  TTGTATCTGT  ACCCGTGGCA  TCTCCAGAAT
5601  CTAAACAAAG  TCGAGAGCCG  GAACCGGCTG  CTGTTCCAGC  ATCAGTTTCT
5651  GTTAGAGAGA  TGTGGAGCAC  GGCTGACAGA  TTGACCACCG  GTACATGCCC
5701  ATCGACTCGA  TGCGGAGCAA  CTAGCTGGGT  AACAGATGGA  ACTAAAGTAA
5751  CAGTTTATGA  AGAAAAAGAC  GGTTGGTCTA  GAATCGGAGA  GCTACAGTCT
5801  GCAATGTGCA  TAAATGGAAT  AAGTGGCGCG  GTCGATTCAG  GTGAATCTTC
5851  CTGCAATCCC  ACCAATGGTA  TCGTTAATGG  GCAATTCGCA  CCCTGGGTTT
5901  TCTCGGATTA  TCTTACGATC  CAAGAGCCAG  AAGCTCCCAT  ATCCACCCAA
5951  GAGTGTCGAA  ATATGGGGCT  CGAGAACTCA  GATAATTACC  GTATCTATTC
6001  TAGTCAGTTC  TGCACTGCCG  CTCTCGAAAT  GATCAACGAT  AGAGTATGCA
6051  ATACATCTGA  TTTCAGAGAT  TTAGCTTGGT  TATCTTCTCC  TGAAAGAGGA
6101  CAGGATTACT  ACTTCACCTA  TTGTGTGGCGA  TTTCAACCTC  AAAACAGATG
6151  GTATTTGAAT  GTCAGGACAG  GTGAAATCAC  CCGCTGATAT  TCCACCAAGG
```

FIGURE 2C

```
6201   TGAGTCCTGT AGATCAGACT CTCAAGGAGT AAACGTTTTA ATCCATCTCC
6251   ATGAGATCAA CATAGATAGG TGTTCAGTCC CGGCATCTGG TGGATCGGGT
6301   TTAGGATGAA TCTGTCCGGC TCTTGACATA CCCCCGCGTG AAACCCTGTC
6351   TTTACAAGAG AAAGTCAGCG GCCTCGAAGC CGCTCTAGCC GATGCCCGGG
6401   CCCAACGGGA TGAGTAGAGC GAACAAGCAA AGCGCCTAGC TATGGCTCTG
6451   CCCGTCCCGG AAGCTGCAGC CGCAGAATCC GGAAAAAAGA AAAAATACAT
6501   GGCACGCGATT ATTTGGATAG GACACAATCC TTTTCTATTA ATATACAACA
6551   AGATATGGGC ATGCGCCGCG CGTGATCCTC ATTCGATACA ATCCAAATCC
6601   TGAAAGCTGA CTATGCCCTA CGCATCGCGC ACCATCGGTG CCGTCATTGA
6651   TGACGTGAAC CGCACCTACC TGCTGCCCGC AATCCAACGC CCCTATGTCT
6701   GGTCTGCCGG ACAGGTCGTT GCGCTGTTCG ACTCTCTGTT GAAGGGCTAT
6751   CCGATCAGCA GCTTCATGTT CTGGGCGGTG GACGAGGAGA CCAAGGCAGA
6801   GCTGCGATGC TACAAATTCA TCGAGAATTA TCGGCCCGAA ATGATGAACG
6851   AGCCGACTAG TGCGGACGGG CGGCAGGTCG TCCTTGTGCT CGACGGACAG
6901   CAGCGGATGA CCTCACTGTT GATCGGCTTG CGCGGCACAT TCTCTGAGAA
6951   AGCCAAACAC GCGCGCAACA GCAACGCGGC CGCGTGGTCG GCAAAAACGC
7001   TATATCTAGA CCTGCTTCGG GACCCGGATC CGAAGAACTC CGATGAAGAC
7051   GAAGGCAATG AGTTCGGAAT CACTTACGGT CTCTCTTTCC ATGAACGCCG
7101   CCCGACCAGC AGCCACAGGC ACCACTGGTT CAAGGTGGGA TCGATACTGG
7151   ATTATCCTAC AGACGAGCAG CTGGAGGGGT TGATTGCCAA GGTGAAGACC
7201   GAATTTCATC ATGGTGTATC GGATTGGGAA AAGGGGCTGG CGGAAGACAC
7251   CCTGCGCCGG TTGCACCGCG TCATCTGGAA AGACGAGGGC ATCAACTTTT
7301   TCACTGAACG CGACCAGTCG GTTGATCGGG TGCTGGACAT CTTCGTGCGG
7351   GCCAATGACG GGGGCACGAA ACTGTCGAAG GCAGACCTGC TGATGTCGAT
7401   GATCACGTCA AAATGGTCCA GCGGATCGGC CCGCGAGGAA ATCGGCGGCT
7451   TTGTCGAGCA CATAAACAAA GGTCTCGGTG CGCCGAACAA GATCAGTCGC
7501   GATCTGGTCC TGAAGGCCTG TCTGGTCGTC TGCGATTATG ATGTCGTCTA
7551   TAATGTCAGG AACTTTACAA GCGAGGTCAT CGGCAGGATC GAAAGCAACT
7601   GGGATCGTAT CAAGCAGGCA TTCGAGAACA CGTTCCGCCT GCTGAACAGG
7651   CATGGCATCA CCGGGGATAA CCTCGGCTCT TTGAACGCGG TGCTGCCTCT
7701   GGTCTATTAT ATCTACAACA CGCCGGATTT CGATTTCCGA GGATCGAGCG
7751   AGTTCGAGCG GGTCAATGCC AGCTCCATGC ACCTCTGGTT GGTGAACAGC
7801   CTGCTGGTCA GCGCCTTCGT TGGCCATTCG GATCAGACCA TCACCACCGC
7851   GCGCAATACG ATCCGCGATC ACCTGCGTGT AGGCCGCGAT TTCCCAGTAC
7901   GAAAGCTGTT CGATGCCATG GCGAAGGGGG GACGGCTATC TCAGGTGGAT
7951   GAGCGTACCA TCGAAGAATT GCTGGAAATG CAATATGGCA AGCCCCGGAC
8001   CTTCGTTGCG CTGTCGCTGC TCTATCAGGG CATCGACTGG AACGGATCGA
8051   CCTGGCATGT CGATCATATC ATTCCCCAAG CGGACGCTCA GAAAAATGTG
8101   CTGCGCGGGC GCAATCTGCC CGAGCATCGC ATTCAGGAAA TCTTGGGCGC
8151   GGTTAACAGT TTGGGGCAACC TGCAACTTTT GCGCGGAGAT GAGAATATCG
8201   AGAAAGGTGC GCTGCCATTC AGGTCATGGA TTACCGGACG GCGCGTTGAT
8251   TTCTACGAGC AGCATATGAT CCCGGCGCAC CTTGAACTGT GCGATGTACT
```

FIGURE 2D

```
8301   GCATCTGCCC GAGTTCGTGC GCGAACGGGA AAAGGTGATC CGGCGCCGTT
8351   TGATGGAGTT GGTCGGAGCA CGACGCGCAT GAATGAGGTC GTCTTGTCAC
8401   GCGAAGAGCT GCGTCAATCT TGTCTCGACC TGCTTGAAAA ACGCGCTGTC
8451   GAACACCCTG CGGGACACCA AGGCAAGCTC GCCGCCCGCT ATGTTGTGCA
8501    CCGCGACGA
```

FIGURE 2E

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
 101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC ATTCGCCATT
 251  CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT
 301  TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA
 351  ACGCCAGGGT TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT
 401  CGGCAATGGG TCGAAATTCA TAGAATTTTG TGTGAGGTGC GTAGCGGCTC
 451  TGACAGGGGT GCTGCGCGGA GATCTCTGGT CTCAGGTAGG GCGACAATGG
 501  AGAGGTGTTA GTTGCCCCCT GTATCGCTCT CTGCGTGGCG CATTGGGTCA
 551  TCCTGCCCGG ACATATGATA TTCCGCTAGA GGATTACTGA TAGTTTCTGC
 601  CTGTCGGGCT TGTCGGGCTT GTCGGGCTTG TCGGGCTTGT CGGGCCTGTC
 651  CCTCTTGTCC CGCCTGTCCT CACTTTTTCA CAATCAAAAA ATGGGCGAAG
 701  CCCTTCTTGT TCTATAGTTC TTATAGTTCA TACGAAAATT ACACATAATT
 751  ATCAATAGCT TATTCGCTTA AAAGGGAGTA ATTGGGCCGC AAAAGGGAGT
 801  AATTGGGCCG CAAAAGGGAG TAATTGGGCC GCAAAAGGGA GTAATTGGGC
 851  CGATATCGGT TGTTTACATG GGGAGGAATC CCCTTAATCA TTTCTCCCCA
 901  TGGGAAAGAC AACACAAGTG GCCGCAGACC GGGCCTTCGA CCAGACAAAA
 951  ACTGTGCTCC CTGCCGAGGT GGCGAGAGGG GTCTATATGC GCAATCCGCC
1001  CCGCCTGCAG GCGCTCAAGC TCATGCATTT AATGATAGCC ACTGCGGGCG
1051  GCCGCATGGC TGATGATGTG CGCCATGAAA TGCGGCTGGC CGACATTCGC
1101  GCAATCGACG GCATGAAAAA CCATGACCGT GAGAGCCTGA CCCCGCTGTT
1151  CGAGGAGCTA GCCGCTGCGG TGTTGACCCA TGATGACCCT GCAAAGATGA
1201  TCGTGACAGT CGGCGGCTTG GTCGATGAGG CGCGAATAGA CTACCGCCAG
1251  GAGGCAAGCG GCGAACTCCT AGTGACGTGG ACCTTCCGGA GTACATTCCG
1301  TCGTATGGCG GCGGAGTCGA ACCACTGGGC CATTCTCGAC CGTCAAACGG
1351  TATTCCATCT CGGTAGTAAG TATTCCGTGC TGCTGTTCCA GCACGTCTCT
1401  AGTCTCGCCA ATCTTGATCG GATGAGCGCG AAAACCTTTA CGGTCCCCGA
1451  GTTGCGGGCG CTCCTTGGAG TGCCCGAGGG AAAGATGGTT CGTTGGAACG
1501  ACGTTAACAG ATTTGCTCTC AAACCTGCAC TGGATGAGAT CAACCATTTA
1551  TCGCGTCTGA CATTGACGGC AAAGCCGACC AAGATTGGCC GTAGCGTGGC
1601  AAGTGTGACT ATAGGCTGGG AAGTGAAAGA CGACCCAACC GTCGCCAGGC
1651  GCGAGCTGGC GGGTTCCAAG GTCGGTCGAG ATGCTCGTCG CAGAGGGGCA
1701  GCGGAAACGA TAGCCCCCTC CTTCCCAGAA GCGGGCGGGA TCACCTACAG
1751  TCCACGTTGG CTGGAGCTGA AACGCTCTGC TGGCAGCAAC AAGGACAACG
1801  ATCTGATCGC CTCAGACTTC CGGCGTTTCT GTCGGGAGAG AGGCGTGCGT
1851  CTGGACGCTG CAAACATCGA AAAACTGTTT TTAGATTTCT GCGCAAAGGT
1901  AGGGAAGGTT TGAGTTTGTA GGTATTTCAC CGCAATAGTG TTAAATGACT
1951  TTCGTGAAAC GATGTGCAAT ATAGCGGTAA GACTATGAAA TACACGGCTG
```

FIGURE 3A

```
2001   GACAGGCTGC AAAAGCAACG GGTGTGGCGA CCGCAACCAT CACTCGGGCG
2051   CTAAAAAGCG GTAAAATTTC CGGTAAAAAA GATGAATCTG GGGCATGGGT
2101   TATAGATCCT GCAGAATTGC ACAGAGTGTT TCCTCCCATT TCAAAGAAAT
2151   ACACCGAAAC ACCTAACACG CAAGTATATG GTAAGCGTGA TGAAACACAT
2201   GAAATGACCT CAGAAATCAG CGCATTAGAG CGTGAAGTTC GGACTTTACG
2251   CGATGCTTTA TCTGATGCCA GGGAGGATCG CGACAAATGG CGCGACATGG
2301   CCGAGCGTCT TTCAATTTCA TCACCGATGA GAGAGGAAGA CCGCCCCCCT
2351   CAAAAACAAA GATGGTGGAA GATATTCTGA TCCTGGGCTT CAGGAGCCTT
2401   GCCTTTACTG GCGGAAAAAC GCGATATTGA GGCACAGGCC CGCACTTTAG
2451   AGGCGGAAGC CTATAACGAG TACCAAAACA CTAGAAGCCA GATTGAGGAA
2501   AATAGGGAAC GTGGGATCCT CTAGAGTCGA CCTGCAGGCA TGCAAGCTTG
2551   GCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC
2601   AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG
2651   CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT
2701   TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAACG
2751   CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CTCTTCCGCG CTCGGTCTTG
2801   CCTTGCTCGT CGGTGATGTA CTTCACCAGC TCCGCGAAGT CGCTCTTCTT
2851   GATGGAGCGC ATGGGGACGT GCTTGGCAAT CACGCGCACC CCCCGGCCGT
2901   TTTAGCGGCT AAAAAAGTCA TGGCTCTGCC CTCGGGCGGA CCACGCCCAT
2951   CATGACCTTG CCAAGCTCGT CCTGCTTCTC TTCGATCTTC GCCAGCAGGG
3001   CGAGGATCGT GGCATCACCG AACCGCGCCG TGCGCGGGTC GTCGGTGAGC
3051   CAGAGTTTCA GCAGGCCGCC CAGGCGGCCG AGGTCGCCAT TGATGCGGGC
3101   CAGCTCGCGG ACGTGCTCAT AGTCCACGAC GCCCGTGATT TTGTAGCCCT
3151   GGCCGACGGC CAGCAGGTAG GCCGACAGGC TCATGCCGGC CGCGCCGCC
3201   TTTTCCTCAA TCGCTCTTCG TTCGTCTGGA AGGCAGTACA CCTTGATAGG
3251   TGGGCTGCCC TTCCTGGTTG GCTTGGTTTC ATCAGCCATC CGCTTGCCCT
3301   CATCTGTTAC GCCGGCGGTA GCCGGCCAGC CTCGCAGAGC AGGATTCCCG
3351   TTGAGCACCG CCAGGTGCGA ATAAGGGACA GTGAAGAAGG AACACCCGCT
3401   CGCGGGTGGG CCTACTTCAC CTATCCTGCC CGGCTGACGC CGTTGGATAC
3451   ACCAAGGAAA GTCTACACGA ACCCTTTGGC AAAATCCTGT ATATCGTGCG
3501   AAAAAGGATG GATATACCGA AAAAATCGCT ATAATGACCC CGAAGCAGGG
3551   TTATGCAGCG GAAAAGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCGT
3601   TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT
3651   CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
3701   CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
3751   GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA AGTCAGAGGT
3801   GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC
3851   TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
3901   CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA
3951   GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC
4001   GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
4051   TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
```

FIGURE 3B

```
4101   GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG
4151   AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT TTGGTATCTG
4201   CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT
4251   CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG
4301   CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTTTC
4351   TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG
4401   TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT GGGGGGGGG
4451   GGCGCTGAGG TCTGCCTCGT GAAGAAGGTG TTGCTGACTC ATACCAGGCC
4501   TGAATCGCCC CATCATCCAG CCAGAAAGTG AGGGAGCCAC GGTTGATGAG
4551   AGCTTTGTTG TAGGTGGACC AGTTGGTGAT TTTGAACTTT TGCTTTGCCA
4601   CGGAACGGTC TGCGTTGTCG GGAAGATGCG TGATCTGATC CTTCAACTCA
4651   GCAAAAGTTC GATTTATTCA ACAAAGCCGC CGTCCCGTCA AGTCAGCGTA
4701   ATGCTCTGCC AGTGTTACAA CCAATTAACC AATTCTGATT AGAAAAACTC
4751   ATCGAGCATC AAATGAAACT GCAATTTATT CATATCAGGA TTATCAATAC
4801   CATATTTTTG AAAAAGCCGT TTCTGTAATG AAGGAGAAAA CTCACCGAGG
4851   CAGTTCCATA GGATGGCAAG ATCCTGGTAT CGGTCTGCGA TTCCGACTCG
4901   TCCAACATCA ATACAACCTA TTAATTTCCC CTCGTCAAAA ATAAGGTTAT
4951   CAAGTGAGAA ATCACCATGA GTGACGACTG AATCCGGTGA GAATGGCAAA
5001   AGCTTATGCA TTTCTTTCCA GACTTGTTCA ACAGGCCAGC CATTACGCTC
5051   GTCATCAAAA TCACTCGCAT CAACCAAACC GTTATTCATT CGTGATTGCG
5101   CCTGAGCGAG ACGAAATACG CGATCGCTGT TAAAAGGACA ATTACAAACA
5151   GGAATCGAAT GCAACCGGCG CAGGAACACT GCCAGCGCAT CAACAATATT
5201   TTCACCTGAA TCAGGATATT CTTCTAATAC CTGGAATGCT GTTTTCCCGG
5251   GGATCGCAGT GGTGAGTAAC CATGCATCAT CAGGAGTACG GATAAAATGC
5301   TTGATGGTCG GAAGAGGCAT AAATTCCGTC AGCCAGTTTA GTCTGACCAT
5351   CTCATCTGTA ACATCATTGG CAACGCTACC TTTGCCATGT TTCAGAAACA
5401   ACTCTGGCGC ATCGGGCTTC CCATACAATC GATAGATTGT CGCACCTGAT
5451   TGCCCGACAT TATCGCGAGC CCATTTATAC CCATATAAAT CAGCATCCAT
5501   GTTGGAATTT AATCGCGGCC TCGAGCAAGA CGTTTCCCGT TGAATATGGC
5551   TCATAACACC CCTTGTATTA CTGTTTATGT AAGCAGACAG TTTTATTGTT
5601   CATGATGATA TATTTTTATC TTGTGCAATG TAACATCAGA GATTTTGAGA
5651   CACAACGTGG CTTTCCCCCC CCCCCCATTA TTGAAGCATT TATCAGGGTT
5701   ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA
5751   ATAGGGGTTC CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA
5801   AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT ATCACGAGGC
5851   CCTTTCGTC
```

FIGURE 3C

```
   1  GGCAATGGGT CGAAATTCAT AGAATTTGT GTGAGGTGCG TAGCGGCTCT
  51  GACAGGGGTG CTGCGCGGAG ATCTCTGGTC TCAGGTAGGG CGACAATGGA
 101  GAGGTGTTAG TTGCCCCCTG TATCGCTCTC TGCGTGGCGC ATTGGGTCAT
 151  CCTGCCCGGA CATATGATAT TCCGCTAGAG GATTACTGAT AGTTTCTGCC
 201  TGTCGGGCTT GTCGGGCTTG TCGGGCTTGT CGGGCTTGTC GGGCCTGTCC
 251  CTCTTGTCCC GCCTGTCCTC ACTTTTTCAC AATCAAAAAA TGGGCGAAGC
 301  CCTTCTTGTT CTATAGTTCT TATAGTTCAT ACGAAAATTA CACATAATTA
 351  TCAATAGCTT ATTCGCTTAA AAGGGAGTAA TTGGGCCGCA AAAGGGAGTA
 401  ATTGGGCCGC AAAAGGGAGT AATTGGGCCG CAAAAGGGAG TAATTGGGCC
 451  GATATCGGTT GTTTACATGG GGAGGAATCC CCTTAATCAT TTCTCCCCAT
 501  GGGAAAGACA ACACAAGTGG CCGCGACACCG GGCCTTCGAC CAGACAAAAA
 551  CTGTGCTCCC TGCCGAGGTG GCGAGAGGGG TCTATATGCG CAATCCGCCC
 601  CGCCTGCAGG CGCTCAAGCT CATGCATTTA ATGATAGCCA CTGCGGGCGG
 651  CCGCATGGCT GATGATGTGC GCCATGAAAT GCGGCTGGCC GACATTCGCG
 701  CAATCGACGG CATGAAAAAC CATGACCGTG AGAGCCTGAC CCCGCTGTTC
 751  GAGGAGCTAG CCGCTGCGGT GTTGACCCAT GATGACCCTG CAAAGATGAT
 801  CGTGACAGTC GGCGGCTTGG TCGATGAGGC GCGAATAGAC TACCGCCAGG
 851  AGGCAAGCGG CGAACTCCTA GTGCAGTGGA CCTTCCGGAG TACATTCCGT
 901  CGTATGGCGG CGGAGTCGAA CCACTGGGCC ATTCTCGACC GTCAAACGGT
 951  ATTCCATCTC GGTAGTAAGT ATTCCGTGCT GCTGTTCCAG CACGTCTCTA
1001  GTCTCGCCAA TCTTGATCGG ATGAGCGCGA AAACCTTTAC GGTCCCCGAG
1051  TTGCGGGCGC TCCTTGGAGT GCCCGAGGGA AAGATGGTTC GTTGGAACGA
1101  CGTTAACAGA TTTGCTCTCA AACCTGCACT GGATGAGATC AACCATTTAT
1151  CGCGTCTGAC ATTGACGGCA AAGCCGACCA AGATTGGCCG TAGCGTGGCA
1201  AGTGTGACTA TAGGCTGGGA AGTGAAAGAC GACCCAACCG TCGCCAGGCG
1251  CGAGCTGGCG GGTTCCAAGG TCGGTCGAGA TGCTCGTCGC AGAGGGGCAG
1301  CGGAAACGAT AGCCCCCTCC TTCCCAGAAG CGGGCGGGAT CACCTACAGT
1351  CCACGTTGGC TGGAGCTGAA ACGCTCTGCT GGCAGCAACA AGGACAACGA
1401  TCTGATCGCC TCAGACTTCC GGCGTTTCTG TCGGGAGAGA GGCGTGCGTC
1451  TGGACGCTGC AAACATCGAA AAACTGTTTT TAGATTTCTG CGCAAAGGTA
1501  GGGAAGGTTT GAGTTTTGAG GTATTTCACC GCAATAGTGT TAAATGACTT
1551  TCGTGAAACG ATGTGCAATA TAGCGGTAAG ACTATGAAAT ACACGGCTGG
1601  ACAGGCTGCA AAAGCAACGG GTGTGGCGAC CGCAACCATC ACTCGGGCGC
1651  TAAAAAGCGG TAAAATTTCC GGTAAAAAAG ATGAATCTGG GGCATGGGTT
1701  ATAGATCCTG CAGAATTGCA CAGAGTGTTT CCTCCCATTT CAAAGAAATA
1751  CACCGAAACA CCTAACACGC AAGTATATGG TAAGCGTGAT GAAACACATG
1801  AAATGACCTC AGAAATCAGC GCATTAGAGC GTGAAGTTCG GACTTTACGC
1851  GATGCTTTAT CTGATGCCAG GGAGGATCGC GACAAATGGC GCGACATGGC
1901  CGAGCGTCTT TCAATTTCAT CACCGATGAG AGAGGAAGAC CGCCCCCCTC
1951  AAAAACAAAG ATGGTGGAAG ATATTCTGAT CCTGGGCTTC AGGAGCCTTG
```

FIGURE 4A

```
2001  CCTTTAAAAC CTGAATCAGC ATTCTAGCGA TGCTGATAAG AAGTAAATAT
2051  AGCCACAATA GAGCGGCCAT TTTCCATTCA CATACAGCTC ATCATGTGAT
2101  CAATATCAAG TATTGATATT CATCAATGGA GAAGAATTTA CATGTATCAC
2151  AGGATCATCA CAGCATTTGT TTTTGTATTT CTAAGTGCTA ACATAACTAT
2201  CGCTGGCCCT AAAGAAGATT GTACTATTGC AGTATCTCAC CTTGGGTTTC
2251  AGACCGATAA TTACAGCTTT GTCGAAGCCG GTTTTTTTGC CAGAGAGAGA
2301  CACGTTTTTG ATGGTGTAAT AAACTGCTAC GTATCTCATG ATGGTAACAT
2351  ACACAGCATC ATCCGGGGCA ACACACCTCT TATGGAAGAT GGATATTATG
2401  GCCCAGAAGT ACTGGCGGAA AAACGCGATA TTGAGGCACA GGCCCGCACT
2451  TTAGAGGCGG AAGCCTATAA CGAGTACCAA AACACTAGAA GCCAGATTGA
2501  GGAAAATAGG GAACGTG
```

FIGURE 4B